

0/20



OIPE

## RAW SEQUENCE LISTING

DATE: 08/28/2003

PATENT APPLICATION: US/10/642,531

TIME: 14:30:32

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\08282003\J642531.raw

3 <110> APPLICANT: Marillia, Elizabeth-France  
 4 Zou, Jitao  
 5 Taylor, David C.  
 7 <120> TITLE OF INVENTION: PLANT PYRUVATE DEHYDROGENASE KINASE GENE  
 9 <130> FILE REFERENCE: 3015-5442US  
 C--> 11 <140> CURRENT APPLICATION NUMBER: US/10/642,531  
 12 <141> CURRENT FILING DATE: 2003-08-15  
 14 <150> PRIOR APPLICATION NUMBER: 10/222,075  
 15 <151> PRIOR FILING DATE: 2002-08-16  
 17 <160> NUMBER OF SEQ ID NOS: 19  
 19 <170> SOFTWARE: PatentIn version 3.1  
 22 <210> SEQ ID NO: 1  
 23 <211> LENGTH: 1104  
 24 <212> TYPE: DNA  
 25 <213> ORGANISM: Brassica napus  
 27 <220> FEATURE:  
 28 <223> OTHER INFORMATION: PDHK cDNA from B. napus  
 30 <400> SEQUENCE: 1  
 31 atggcggtga agaaggctag cgagatgttt tcgaagagct tgatcgagga cgttcacaga 60  
 33 tgggggatgca tgaagcagac gggcgtgagc ctcaggtaca tgatggagtt cggttccact 120  
 35 cccactgaga gaaaccttct gatctcggcg cagtttcttc acaaggagct tccgattcgg 180  
 37 atcgcgaggc gtgcgatcga actcgagacg ctgccttatg gcctctctga gaaacctgcc 240  
 39 gtcttgaagg taagagattg gtatgtggag tcattcaggg acatgagagc gtttctctgag 300  
 41 atcaaggata ctgctgatga gaaagagttc acacagatga tcaaggctgt taaagtaagg 360  
 43 cacaacaacg tggttcccat gatggctctg ggtgtgaacc agctgaagaa aggaatgaaa 420  
 45 ctctacgaaa agcttgatga gattcatcag tttcttgatc gcttctactt gtctcgtata 480  
 47 gggatccgta tgcttatcgg gcagcatggt gagttgcata atccaaaccc accacttcac 540  
 49 acagtgggtt acatacacac caagatgtct cctatggagg tggcaaggaa tgctagttaa 600  
 51 gatgcaaggt cgatttggtt cagagagtat ggttctgctc cggagataaa catatatggc 660  
 53 gatccaagtt tcacttttcc gtatgttccg acccatttgc atcttatggt gtatgagtta 720  
 55 gtcaagaact ctctccgtgc tgtccaagag cggtttggtg actctgatag ggttgaccca 780  
 57 ccaatccgta tcattgttgc tgatggaatc gaagatgtta caataaaggc ctcatatgaa 840  
 59 ggtggaggta taccgagaag cgggtctccct aaaatattca cttacctcta cagcactgca 900  
 61 agaaaccac ttgaagaaga tgtggacttg ggaaccgctg atgttcccct gactatggct 960  
 63 ggttatgggt atgggtctgcc tattagtgcg ttgtatgctc gctattttgg tggagatttg 1020  
 65 cagatcatat ccatggaagg atacgggact gatgcttact tgcacttgct tcgtcttgga 1080  
 67 gactcgcagg agcctttgcc atga 1104  
 70 <210> SEQ ID NO: 2  
 71 <211> LENGTH: 1104  
 72 <212> TYPE: DNA  
 73 <213> ORGANISM: Brassica rapa  
 75 <220> FEATURE:  
 76 <223> OTHER INFORMATION: PDHK cDNA from B. rapa

ENTERED

P.4

## RAW SEQUENCE LISTING

DATE: 08/28/2003

PATENT APPLICATION: US/10/642,531

TIME: 14:30:32

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\08282003\J642531.raw

78 &lt;400&gt; SEQUENCE: 2

```

79 atggcggtga agaaggctag cgagatgttt tcgaagagct tgatcgagga cgttcacaga      60
81 tggggatgca tgaagcagac gggcgtgagc ctcaggtaca tgatggagtt cggttcact      120
83 cccactgaga gaaaccttct gatctcggcg cagtttcttc acaaggagct tccgattcgg      180
85 atcgcgaggc gtgcgatcga actcgagacg ctgccttatg gcctctctga gaaacctgcc      240
87 gtcttgaagg tgagggattg gtatgtggag tcattcaggg acatgagagc gtttcctgag      300
89 atcaaggata ctgctgatga gaaagagttc actcagatga ttaaggctgt taaagtaagg      360
91 cacaacaacg tggttcccat gatggctctg ggtgtgaacc agctgaagaa aggaatgaaa      420
93 ctctacgaaa agcttgatga gattcatcag ttctctgata gcttctactt gtctcgtata      480
95 gggatccgta tgcttatcgg gcagcatgtt gagttgcata atccaaaccc accacttcac      540
97 acagtgggtt acatacacac caagatgtct cctatggagg tggcaaggaa tgctagttaa      600
99 gatgcaaggt cgatttgttt cagagagtat ggttctgctc cggagataaa catatatggc      660
101 gatccaagtt ccactyttcc gtatgttccg acccatttgc atcttatggt gtatgagtta      720
103 gtcaagaact ctctccgtgc tgtccaagag cggtttgttg actctgatag ggttgcacca      780
105 ccaatccgta tcattgttgc tgatggaatc gaagatgtta caataaaggc ctcagatgaa      840
107 ggtggaggtta taccgagaag cggctctccct aaaatattca cttacctcta cagcactgca      900
109 agaaaccacac ttgaagaaga tgtggacttg ggaaccgctg atgttcccct gactatggct      960
111 ggttatggtt atggtctgcc tattagtcgc ttgtatgctc gctattttgg tggagatttg      1020
113 cagatcatat ccatggaagg atacgggact gatgcttact tgcacttgct tcgtcttgga      1080
115 gactcgagcagg agcctttgcc atga                                          1104

```

118 &lt;210&gt; SEQ ID NO: 3

119 &lt;211&gt; LENGTH: 1104

120 &lt;212&gt; TYPE: DNA

121 &lt;213&gt; ORGANISM: Brassica oleracea

123 &lt;220&gt; FEATURE:

124 &lt;223&gt; OTHER INFORMATION: PDHK cDNA from B. oleracea

126 &lt;400&gt; SEQUENCE: 3

```

127 atggcggtga agaaggctag cgagatgttt tcgaagagct tgatcgagga cgttcacaga      60
129 tggggatgca tgaagcagac gggcgtgagc ctcaggtaca tgatggagtt cggttcact      120
131 cccactgaga ggaacctcct gatctcggcg cagtttcttc acaaggagct tccgattcgg      180
133 atcgcgaggc gtgcgatcga actcgagacg ctgccttatg gcctctctga gaaacctgcc      240
135 gtcttgaagg taagrgattg gtatgtggag tcattcaggr acatgagagc gtttcctgag      300
137 atcaaggata ctgctgayga gaaagagttc acacagatga ttaaggctgt taaagtaagg      360
139 cacaacaacg tggttcccat gatggctctg ggtgttaacc agctgaagaa aggaatgaaa      420
141 ctctacgaaa aactcgatga gattcatcag ttctctgata gcttctactt gtcacgtata      480
143 gggatccgta tgcttatcgg gcagcatgtt gagttgcata atccaaaccc accacttcac      540
145 actgtgggtt acatacacac caagatgtct cctatggagg tggcaaggaa tgcyagttaa      600
147 gatgcaaggt cgatttgttt casagagtat ggttctgctc cggagataaa cmtatatggc      660
149 gatccaagtt tcacctttcc gtatgtacca acccatttgc atcttatggt gtatgagcta      720
151 gtcaagaact ctctacgtgc tgtccaagag cgatttgttg attctgatag ggttgcacca      780
153 ccaatccgta tcattgttgc tgatggaatc gaagatgtta caataaaggc ctcagatgaa      840
155 ggtggaggtta taccgagaag cggctctgcc aaaatattca cttacctsta cagcactgca      900
157 agaaaccgcg ttgaagaaga tgtggacttg ggaacagctg atgtaccgct gacwatggct      960
159 ggttatggtt atggtctgcc yattagtcgc ttgtatgctc gatactttgg tggagatttg      1020
161 cagatcatat ccatggaagg atacgggact gatgcttact tgcacttgct tcgtcttgga      1080
163 gactcgcaag agcctttgcc atga                                          1104

```

166 &lt;210&gt; SEQ ID NO: 4

167 &lt;211&gt; LENGTH: 1104

168 &lt;212&gt; TYPE: DNA

## RAW SEQUENCE LISTING

DATE: 08/28/2003

PATENT APPLICATION: US/10/642,531

TIME: 14:30:32

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\08282003\J642531.raw

169 &lt;213&gt; ORGANISM: Brassica carinata

171 &lt;220&gt; FEATURE:

172 &lt;223&gt; OTHER INFORMATION: PDHK cDNA from B. carinata

174 &lt;400&gt; SEQUENCE: 4

```

175 atggcggtga agaaggctag cgagatgttt tcgaagagct tgatcgagga cgttcacaga      60
177 tggggatgca tgaagcagac gggcgtgagc ctcaggtaca tgatggagtt cggttccact    120
179 cccactgaga ggaacctcct gatctcggcg cagtttcttc acaaggagct tccgattcgg    180
181 atcgcgaggc gtgcgatcga actcgagacg ctgccttatg gcctctctga gaaacctgcc    240
183 gtcttgaagg taagagattg gtatgtggag tcattcaggg acatgagagc gtttctctgag   300
185 atcaaggata ctgctgatga gaaagagttc acacagatga ttaaggctgt taaagtaagg   360
187 cacaacaacg tggttcccat gatggctctg ggtgttaacc agctgaagaa aggaatgaaa   420
189 ctctacgaaa aactcgatga gattcatcag ttttttgatc gcttctactt gtcacgtata   480
191 gggatccgta tgcttatcgg gcagcatggt gagttgcata atccaaaccc accacttcac   540
193 actgtggggt acatacacac caagatgtct ccaatggagg tggcaaggaa tgctagttaa   600
195 gatgcaaggt cgatttggtt ccgagagtat ggttctgctc cggagataaa catatatggc   660
197 gatccaagtt tcacctttcc gtatgtacca acccatttgc atcttatggt gtatgagcta   720
199 gtcaagaact ctctacgtgc tgtccaagag cggtttggtg actctgatag ggttgaccca   780
201 ccaatccgta tcattgttgc tgatggaatc gaagatgtta caataaaggc ctcagatgaa   840
203 ggtggaggta taccgagaag cggcctgccc aaaatattca cttacctcta cagcactgca   900
205 agaaacccgc ttgaagaaga tgtggacttg ggaacagctg atgtaccctg gactatggct   960
207 ggttatggtt atggtctgcc tattagtcgc ttgtatgctc gatactttgg tggagatttg  1020
209 cagatcatat ccattggaag atacgggact gatgcttact tgcacttata tcgtcttgga  1080
211 gactcgcagg agcctttgcc atga                                     1104

```

214 &lt;210&gt; SEQ ID NO: 5

215 &lt;211&gt; LENGTH: 367

216 &lt;212&gt; TYPE: PRT

217 &lt;213&gt; ORGANISM: Brassica napus

219 &lt;220&gt; FEATURE:

220 &lt;223&gt; OTHER INFORMATION: Deduced amino acid sequence from B. napus PDHK cDNA (SEQ ID

NO:1)

222 &lt;400&gt; SEQUENCE: 5

```

223 Met Ala Val Lys Lys Ala Ser Glu Met Phe Ser Lys Ser Leu Ile Glu
224 1          5          10          15
226 Asp Val His Arg Trp Gly Cys Met Lys Gln Thr Gly Val Ser Leu Arg
227          20          25          30
230 Tyr Met Met Glu Phe Gly Ser Thr Pro Thr Glu Arg Asn Leu Leu Ile
231          35          40          45
233 Ser Ala Gln Phe Leu His Lys Glu Leu Pro Ile Arg Ile Ala Arg Arg
234          50          55          60
236 Ala Ile Glu Leu Glu Thr Leu Pro Tyr Gly Leu Ser Glu Lys Pro Ala
237 65          70          75          80
239 Val Leu Lys Val Arg Asp Trp Tyr Val Glu Ser Phe Arg Asp Met Arg
240          85          90          95
242 Ala Phe Pro Glu Ile Lys Asp Thr Ala Asp Glu Lys Glu Phe Thr Gln
243          100         105         110
245 Met Ile Lys Ala Val Lys Val Arg His Asn Asn Val Val Pro Met Met
246          115         120         125
248 Ala Leu Gly Val Asn Gln Leu Lys Lys Gly Met Lys Leu Tyr Glu Lys
249          130         135         140
251 Leu Asp Glu Ile His Gln Phe Leu Asp Arg Phe Tyr Leu Ser Arg Ile

```

## RAW SEQUENCE LISTING

DATE: 08/28/2003

PATENT APPLICATION: US/10/642,531

TIME: 14:30:32

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\08282003\J642531.raw

```

252 145          150          155          160
254 Gly Ile Arg Met Leu Ile Gly Gln His Val Glu Leu His Asn Pro Asn
255          165          170          175
257 Pro Pro Leu His Thr Val Gly Tyr Ile His Thr Lys Met Ser Pro Met
258          180          185          190
260 Glu Val Ala Arg Asn Ala Ser Glu Asp Ala Arg Ser Ile Cys Phe Arg
261          195          200          205
263 Glu Tyr Gly Ser Ala Pro Glu Ile Asn Ile Tyr Gly Asp Pro Ser Phe
264          210          215          220
266 Thr Phe Pro Tyr Val Pro Thr His Leu His Leu Met Val Tyr Glu Leu
267 225          230          235          240
269 Val Lys Asn Ser Leu Arg Ala Val Gln Glu Arg Phe Val Asp Ser Asp
270          245          250          255
272 Arg Val Ala Pro Pro Ile Arg Ile Ile Val Ala Asp Gly Ile Glu Asp
273          260          265          270
275 Val Thr Ile Lys Val Ser Asp Glu Gly Gly Gly Ile Pro Arg Ser Gly
276          275          280          285
278 Leu Pro Lys Ile Phe Thr Tyr Leu Tyr Ser Thr Ala Arg Asn Pro Leu
279          290          295          300
281 Glu Glu Asp Val Asp Leu Gly Thr Ala Asp Val Pro Leu Thr Met Ala
282 305          310          315          320
284 Gly Tyr Gly Tyr Gly Leu Pro Ile Ser Arg Leu Tyr Ala Arg Tyr Phe
285          325          330          335
287 Gly Gly Asp Leu Gln Ile Ile Ser Met Glu Gly Tyr Gly Thr Asp Ala
288          340          345          350
290 Tyr Leu His Leu Ser Arg Leu Gly Asp Ser Gln Glu Pro Leu Pro
291          355          360          365

```

294 &lt;210&gt; SEQ ID NO: 6

295 &lt;211&gt; LENGTH: 367

296 &lt;212&gt; TYPE: PRT

297 &lt;213&gt; ORGANISM: Brassica rapa

299 &lt;220&gt; FEATURE:

300 &lt;223&gt; OTHER INFORMATION: Deduced amino acid sequence from B. rapa PDHK cDNA (SEQ ID

NO:2)

302 &lt;400&gt; SEQUENCE: 6

```

303 Met Ala Val Lys Lys Ala Ser Glu Met Phe Ser Lys Ser Leu Ile Glu
304 1          5          10          15
306 Asp Val His Arg Trp Gly Cys Met Lys Gln Thr Gly Val Ser Leu Arg
307          20          25          30
309 Tyr Met Met Glu Phe Gly Ser Thr Pro Thr Glu Arg Asn Leu Leu Ile
310          35          40          45
312 Ser Ala Gln Phe Leu His Lys Glu Leu Pro Ile Arg Ile Ala Arg Arg
313          50          55          60
315 Ala Ile Glu Leu Glu Thr Leu Pro Tyr Gly Leu Ser Glu Lys Pro Ala
316 65          70          75          80
318 Val Leu Lys Val Arg Asp Trp Tyr Val Glu Ser Phe Arg Asp Met Arg
319          85          90          95
321 Ala Phe Pro Glu Ile Lys Asp Thr Ala Asp Glu Lys Glu Phe Thr Gln
322          100          105          110
324 Met Ile Lys Ala Val Lys Val Arg His Asn Asn Val Val Pro Met Met

```

## RAW SEQUENCE LISTING

DATE: 08/28/2003

PATENT APPLICATION: US/10/642,531

TIME: 14:30:32

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\08282003\J642531.raw

```

325          115          120          125
327 Ala Leu Gly Val Asn Gln Leu Lys Lys Gly Met Lys Leu Tyr Glu Lys
328      130          135          140
330 Leu Asp Glu Ile His Gln Phe Leu Asp Arg Phe Tyr Leu Ser Arg Ile
331 145          150          155          160
333 Gly Ile Arg Met Leu Ile Gly Gln His Val Glu Leu His Asn Pro Asn
334          165          170          175
336 Pro Pro Leu His Thr Val Gly Tyr Ile His Thr Lys Met Ser Pro Met
337          180          185          190
339 Glu Val Ala Arg Asn Ala Ser Glu Asp Ala Arg Ser Ile Cys Phe Arg
340          195          200          205
342 Glu Tyr Gly Ser Ala Pro Glu Ile Asn Ile Tyr Gly Asp Pro Ser Ser
343          210          215          220
345 Thr Phe Pro Tyr Val Pro Thr His Leu His Leu Met Val Tyr Glu Leu
346 225          230          235          240
348 Val Lys Asn Ser Leu Arg Ala Val Gln Glu Arg Phe Val Asp Ser Asp
349          245          250          255
351 Arg Val Ala Pro Pro Ile Arg Ile Ile Val Ala Asp Gly Ile Glu Asp
352          260          265          270
354 Val Thr Ile Lys Val Ser Asp Glu Gly Gly Gly Ile Pro Arg Ser Gly
355          275          280          285
357 Leu Pro Lys Ile Phe Thr Tyr Leu Tyr Ser Thr Ala Arg Asn Pro Leu
358          290          295          300
360 Glu Glu Asp Val Asp Leu Gly Thr Ala Asp Val Pro Leu Thr Met Ala
361 305          310          315          320
363 Gly Tyr Gly Tyr Gly Leu Pro Ile Ser Arg Leu Tyr Ala Arg Tyr Phe
364          325          330          335
366 Gly Gly Asp Leu Gln Ile Ile Ser Met Glu Gly Tyr Gly Thr Asp Ala
367          340          345          350
369 Tyr Leu His Leu Ser Arg Leu Gly Asp Ser Gln Glu Pro Leu Pro
370          355          360          365

```

373 &lt;210&gt; SEQ ID NO: 7

374 &lt;211&gt; LENGTH: 367

375 &lt;212&gt; TYPE: PRT

376 &lt;213&gt; ORGANISM: Brassica oleracea

378 &lt;220&gt; FEATURE:

379 &lt;221&gt; NAME/KEY: misc\_feature

380 &lt;222&gt; LOCATION: (94)..(94)

381 &lt;223&gt; OTHER INFORMATION: X at position 94 is N or D

383 &lt;220&gt; FEATURE:

384 &lt;223&gt; OTHER INFORMATION: Deduced amino acid sequence from B. oleracea PDHK cDNA (SEQ ID NO:3).

386 &lt;400&gt; SEQUENCE: 7

```

387 Met Ala Val Lys Lys Ala Ser Glu Met Phe Ser Lys Ser Leu Ile Glu
388 1          5          10          15
390 Asp Val His Arg Trp Gly Cys Met Lys Gln Thr Gly Val Ser Leu Arg
391          20          25          30
393 Tyr Met Met Glu Phe Gly Ser Thr Pro Thr Glu Arg Asn Leu Leu Ile
394          35          40          45
396 Ser Ala Gln Phe Leu His Lys Glu Leu Pro Ile Arg Ile Ala Arg Arg

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/642,531

DATE: 08/28/2003  
TIME: 14:30:33

Input Set : A:\SEQUENCE LISTING.txt  
Output Set: N:\CRF4\08282003\J642531.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:7; Xaa Pos. 94  
Seq#:9; N Pos. 15,16  
Seq#:10; N Pos. 13

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:9; Line(s) 542  
Seq#:10; Line(s) 564  
Seq#:11; Line(s) 581  
Seq#:12; Line(s) 593  
Seq#:13; Line(s) 605  
Seq#:14; Line(s) 617  
Seq#:15; Line(s) 629

Invalid <213> Response:

Use of "Artificial" only as "<213> Organism" response is incomplete, per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:9,10,11,12,13,14,15,16,17,18,19

**VERIFICATION SUMMARY**

DATE: 08/28/2003

PATENT APPLICATION: US/10/642,531

TIME: 14:30:33

Input Set : A:\SEQUENCE LISTING.txt

Output Set: N:\CRF4\08282003\J642531.raw

L:11 M:270 C: Current Application Number differs, Replaced Current Application Number  
L:402 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:80  
L:555 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9 after pos.:0  
L:572 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10 after pos.:0